Registering α-Helices and β-Strands Using Backbone C—H...O Interactions

S. Kumar Singh, M. Madan Babu, and P. Balaram*
Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

ABSTRACT  The possible occurrence of a novel helix terminating structural motif in proteins involving a stabilizing short C—H...O interaction has been examined using a dataset of 634 non-homologous protein structures (≤2.0 Å). The search for this motif was prompted by the crystallographic characterization of a novel structural feature in crystals of a synthetic decapeptide in which extension of a Schellman motif led to the formation of a C—H...O hydrogen bond between the T-4 C=H and the T+1 C=O groups, where T is the helix terminator adopting a left handed (αL) conformation. More than 100 such motifs with backbone conformation superposing well with the peptide examples were identified. In several examples, formation of this motif led to an approximately antiparallel arrangement of a helical segment with an extended β-strand. Careful examination of these examples suggested the possibility of registering antiparallel arrangement of helices and strands by means of backbone C—H...O interactions with a regular periodicity. Model building resulted in the generation of idealized αβ and βα motifs, which can then be generalized to higher-order repetitive structures. Inspection of the antiparallel αβ motif revealed a significant propensity for Ser, Glu, and Gln residues at the T-4 position resulting in further stabilization using an O...H—N side-chain—backbone hydrogen bond. Modeling studies revealed ready accommodation of serine residues along the helix face that contacts the strand. The theoretically generated folds correspond to “open” polypeptide structures. Proteins 2003;51:167–171.

© 2003 WILEY-LISS, INC.

Key words: polypeptide folds; C—H...O hydrogen bonds; helix termination; αβ motifs

INTRODUCTION

Variations in the spatial arrangements of α-helices and β-sheets give rise to the diversity of polypeptide folds observed in globular protein structures.1–3 While the individual structural elements, helices, and sheets are stabilized by the cooperative formation of multiple hydrogen bonds,4,5 the precise arrangement of secondary structural elements is generally determined by the tertiary interactions involving amino acid side-chains.6–8 The serendipitous observation of an interesting helix termination motif stabilized by a potential C—H...O hydrogen bond, in the crystal structure of a synthetic peptide,9 prompted us to examine the possibility of organizing structures containing α-helices and β-strands using backbone C—H...O interactions, in order to achieve appropriate registry. This approach permits the visualization of a novel class of polypeptide folds in which approximately antiparallel arrangements of helices and strands may be achieved purely by favorable backbone interactions.

Figure 1 shows the superposition of the helix termination motif observed in the decapeptide9 Boc-LUVALUV-DA-DL-U-OMe with eight examples of the similar motifs observed in protein structures (P=A—d-alanine; P=L—d-leucine; U = α-aminoisobutyric acid, Aib). The superposed examples are: 1MROA, Y432 to L441; 1AXN, K59 to E68; 1E2UA, A100 to D109; 1EW4A, Q93 to T102; 1AXN, D215 to D224; 1HVBA, M169 to S178; 1D5TA, D419 to A428; 1DIXA, N147 to T156. The observed root-mean-square deviation is 0.65 Å (Cα atoms only).

© 2003 WILEY-LISS, INC.
DL-U-OMe (U = Aib, α-aminoisobutyric acid), with eight similar examples obtained from a dataset of 634 high-resolution protein structures. This stereochemical feature is characterized by the termination of an α-helix in a Schellman motif, where the terminating residue “T” adopts a left-handed (L) conformation. A short C—H…O interaction between the T-4 C—H and the T-1 C=O forms a novel feature of the motif (C—H…O ≤ 3.5 Å) when the T-1 residue adopts an extended (β) conformation. The possible importance of C—H…O interactions as a stabilizing feature in crystals was recognized almost four decades ago. The existence of C—H…O hydrogen bonds in collagen was considered in early structural studies, with firm experimental evidence appearing much later. Several recent studies have emphasized the role of weak C—H…O interactions in organic and biological molecules.

RESULTS AND DISCUSSION

A survey of a dataset of 634 non-homologous (≤25% sequence homology) protein structures from the Protein Data Bank (resolution ≤ 2.0 Å) revealed as many as 111 examples of the motif (Babu et al.). In almost half these cases, the α-helix was followed by a β-strand, with the long axis of the two elements making an angle less than 40°. A few examples were identified in which an approximate antiparallel arrangement of helices and strands was achieved, wherein two short C—H…O interactions, T-4 C—H to T+1 C=O and T-8 C—H to T+3 C=O, could be observed. Figure 2(a) illustrates the structure seen for residues 88 to 105 in the protein CYAY (a member of the frataxin family from Escherichia coli, PDB id: 1EW4, 1.4 Å). Inspection of the structural feature reveals that the T-8 to T-4 H…O distance (5.83 Å) in the helix is approximately 3.5 Å.
approximately equal to the O—O distance of 5.82 Å between the T+1 and T+3 C—O groups. Encouraged by this, we considered the possibility of extending the registry observed between a helix and an extended strand to longer stretches of secondary structures. Using idealized starting geometries, extremely mild energy minimization (100 cycles of minimization with the Insight II software, MSI Inc.) was applied to the starting structure with the proviso that the appropriate C–O distances must be ≤3.5 Å. Figure 2(b) illustrates a stereochemically acceptable alignment of the helix and strand structures (satisfying these constraints). No unfavorable short contacts are observed in the model with all residues lying within the allowed regions of the Ramachandran map. A β hairpin could then be readily generated, with the N-terminal strand as the template, using a nucleating type II’ β-turn. This results in an αββ motif.

We then explored the possibility of registering a strand and a helix with the opposite polarity. Extraction of the βα motifs in proteins, in which the connecting loops are limited to two residues, yielded 202 examples from which nine examples could be visually identified where the strand and the helix had a good antiparallel arrangement. In the case of the antigen 85C from Mycobacterium tuberculosis (PDB id: 1DQZ, 1.5 Å), a potentially useful short C–O distance of 3.75 Å was observed between the residues Gly 122 and Gly 127 [Fig. 3(a)]. Using this as a starting model, a stereochemically satisfactory strand-helix motif with registering C–O ... O interactions between residues H-3 C—O and H-1 N was made (residue H is defined as the first residue in the C-terminal helix). The motif could then be extended as described for the αβ segment [Fig. 3(b)]. Notably, the β-strand and the α-helix are linked by a single residue (H-1) adopting a conformation with φ ~ 67° and ψ ~ −99°, which corresponds to the i+1 residue of the type II’ β-turn. The first residue of the helix (φ = −79°, ψ = −29°) acts as the i+2 residue and a stabilizing 4 → 1 hydrogen bond is formed between the H-2 C—O and the H+1 N—H. The H-2 residue, which occupies the terminal position of the β-strand, adopts an almost totally extended conformation (φ ~ −166°, ψ ~ 161°).

Figure 4(a,b) shows sections of the αββ and ββα motifs that are registered exclusively through backbone interactions in a composite αβββ structure. Figure 4(c) is a ribbon representation of this fold which constitutes an “open” well-ordered structure. Recent examples of structure determination of classes of RNA-binding proteins have revealed structures with open motifs, comprising exclusively small
Examples of open β-sheets have also been reported.

During the course of our analysis of the helix termination motif illustrated in Figure 1, we observed that in the examples extracted from the protein database, the residues serine, glutamic acid, and glutamine showed a significant propensity to occur at the T-4 position of the helix. Inspection of all these examples revealed that an additional side-chain–backbone hydrogen bond between the Oγ of serine or Oε of glutamic acid/glutamine to the backbone NH of the T+3 residue (N…O distance ≤ 3.5 Å) was observed. Whereas the longer side-chain of glutamic acid or glutamine pushes the β-strand away from the helix, the shorter side-chain of serine can be comfortably accommodated with a strong O…H—N hydrogen bond. Computer modeling studies confirmed that serine could be comfortably accommodated into positions T-4, T-8, and T-12 of the N-terminal helix and at positions H+2, H+6, and H+10 of the C-terminal helix. In the motif generated in Figure 4(c), the required backbone conformations all lie
within the region normally populated by L-amino acids, with only the i + 1 position of the type II β-turn requiring a positive ψ value, which in proteins is usually achieved by positioning Gly or Asn residues. On the helix faces, which contact the strands, Gly/Ala/Ser residues can be accommodated in an almost perfectly antiparallel arrangement, whereas moderate-sized side-chains can be inserted with some distortions. Although weak backbone–backbone C–H···O interactions may be used to register helices with an isolated β-strand segment, it is important to note that the N–H groups on the strand backbone will not be hydrogen bonded and inaccessible to solvent, a potentially unfavorable situation. Locating serine residues not be hydrogen bonded and inaccessible to solvent, a

To note that the N

grams.

Academy of Science summer research fellowship pro-

Institute of Science Young fellowship and the Indian

India. M.M.B. acknowledges support from the Indian

Scientific and Industrial Research (CSIR), Government of

receipt of a senior research fellowship from the Council of

dimensions in polypeptide architecture.

observed in peptides and proteins might provide new

new folds by piecing together persistent structural motifs

arrangements that are distinct from the many αβ motifs observed thus far in the structures of globular proteins. Registry of α-helices and β-strands by means of backbone–backbone interactions, sometimes supplemented by stronger side-chain–backbone hydrogen bonds, may provide a means of creating new polypeptide folds. The invention of new folds by piecing together persistent structural motifs observed in peptides and proteins might provide new dimensions in polypeptide architecture.

ACKNOWLEDGMENTS

This approach of connecting α-helices and β-strands by means of only one or two connecting residues leads to arrangements that are distinct from the many αβ motifs observed thus far in the structures of globular proteins. Registry of α-helices and β-strands by means of backbone–backbone interactions, sometimes supplemented by stronger side-chain–backbone hydrogen bonds, may provide a means of creating new polypeptide folds. The invention of new folds by piecing together persistent structural motifs observed in peptides and proteins might provide new dimensions in polypeptide architecture.

REFERENCES

32. Wilmot CM, Thornton JM. Analysis and prediction of the different stereochemical states of β-strands by means of backbone–backbone hydrogen bonds, may provide a means of creating new polypeptide folds. The invention of new folds by piecing together persistent structural motifs observed in peptides and proteins might provide new dimensions in polypeptide architecture.

ACKNOWLEDGMENTS

This work was funded by the Department of Biotecnology (Government of India) as program support to the area of “Drug and Molecular Design.” S.K.S. acknowledges receipt of a senior research fellowship from the Council of Scientific and Industrial Research (CSIR), Government of India. M.M.B. acknowledges support from the Indian Institute of Science Young fellowship and the Indian Academy of Science summer research fellowship programs.

REFERENCES